

LEVEL II

11

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS  
BEFORE COMPLETING FORM

1. REPORT NUMBER 17127.4-A-EL		2. GOVT ACCESSION NO. AD-A110757	3. RECIPIENT'S CATALOG NUMBER 757
4. TITLE (and Subtitle) Recovering Intrinsic Scene Characteristics from Images		5. TYPE OF REPORT & PERIOD COVERED Technical: 28 Sep 80 - 27 Sep 81	
7. AUTHOR(s) Martin A. Fischler Andrew P. Witkin		8. CONTRACT OR GRANT NUMBER(s) DAAG29 79 C 0216	
9. PERFORMING ORGANIZATION NAME AND ADDRESS SRI International Menlo Park, CA 94025		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE Nov 81	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 15	
		15. SECURITY CLASS. (of this report) Unclassified	
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) NA			
18. SUPPLEMENTARY NOTES The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) images vision information processing surfaces perception			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The central theme of our research is the recovery of information about the three-dimensional structure and physical characteristics of surfaces depicted in an image -- their shapes, locations, and photometric properties. The main obstacle to surface recovery is the confounding of the desired properties in the sensory data: images are inherently ambiguous. Our approach to resolving this ambiguity rests on the application of generic, low-level knowledge (e.g., such basic assumptions as surface continuity and general position) to constrain			

DTIC  
ELECTE  
FEB 10 1982

E

AD A110757

DTIC FILE COPY

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

82 02 08 001

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

410381

20. ABSTRACT CONTINUED

the interpretation. The problem may be viewed as that of decomposing the image into its physically meaningful constituents -- surface orientation, reflectance, illumination, and so on. The "intrinsic image model" provides a conceptual and computational framework in which this view is made explicit.

Surface perception plays a fundamental role in early visual processing, both in humans and in machines.

Work on surface perception has focused on the discrimination of edge types (e.g., extremal boundary or cast shadow), on the three-dimensional interpretation of edges, and on surface reconstruction by interpolating from edges and using texture geometry.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



# SRI International



## RECOVERING INTRINSIC SCENE CHARACTERISTICS FROM IMAGES

Interim Technical Report  
Covering the Period 28 September 1980  
to 27 September 1981

SRI Project 1019  
Contract No. DAAG29-79-C-0216

November 1981

Contract Amount: \$374,693  
Effective Date: September 28, 1979  
Expiration Date: September 27, 1982

By: Martin A. Fischier  
Program Director  
(415) 859-5106

Andrew P. Witkin  
Computer Scientist

Artificial Intelligence Center  
Computer Science and Technology Division

### Prepared for:

U. S. Army Research Office  
P. O. Box 12211  
Research Triangle Park, NC 27709

Attn: Information Processing Office

### Sponsored by:

Defense Advanced Research Projects Agency (DoD)  
ARPA Order No. 2894  
Monitored by U. S. Army Research Office  
Under Contract No. DAAG29-79-C-0216

81 12 28 084

## CONTENTS

LIST OF ILLUSTRATIONS	. . . . .	iii
I INTRODUCTION	. . . . .	1
II RESEARCH ACCOMPLISHMENTS	. . . . .	2
A. Edge Classification	. . . . .	2
B. Defining the Problem	. . . . .	3
C. Computational Theory	. . . . .	5
D. Implementation and Results	. . . . .	7
REFERENCES	. . . . .	9

## ILLUSTRATIONS

1	IDEALIZED REGRESSION PLOTS . . . . .	10
2	EXAMPLE OF EXTREMAL EDGE . . . . .	11
3	EXAMPLE OF EXTREMAL EDGE . . . . .	12
4	EXAMPLE OF LOW-CONTRAST EXTREMAL EDGE . . . . .	13
5	EXAMPLE OF CAST-SHADOW EDGE . . . . .	14
6	EXAMPLE OF REGRESSION PLOTS WHERE NO EDGE IS PRESENT . . . . .	15

## I INTRODUCTION

The central theme of our research is the recovery of information about the three-dimensional structure and physical characteristics of surfaces depicted in an image -- their shapes, locations, and photometric properties. The main obstacle to surface recovery is the confounding of the desired properties in the sensory data: images are inherently ambiguous. Our approach to resolving this ambiguity rests on the application of generic, low-level knowledge (e.g., such basic assumptions as surface continuity and general position) to constrain the interpretation. The problem may be viewed as that of decomposing the image into its physically meaningful constituents -- surface orientation, reflectance, illumination, and so on. The "Intrinsic image model" provides a conceptual and computational framework in which this view is made explicit.

Surface perception plays a fundamental role in early visual processing, both in humans and in machines. An explicit representation of surface structure is necessary for many low-level visual functions involved in such applications as terrain modeling, remote sensing, navigation, manipulation, and obstacle avoidance. It is also a prerequisite for general-purpose vision systems capable of human-level performance in such tasks as object recognition and scene description.

Work on surface perception has focused on the discrimination of edge types (e.g., extremal boundary or cast shadow), on the three-dimensional interpretation of edges, and on surface reconstruction by interpolating from edges and using texture geometry.

## ii RESEARCH ACCOMPLISHMENTS

Much of our earlier work on three-dimensional interpretation of edges and texture assumed a capability to discriminate edges of distinct physical types: extremal edges, shadow boundaries, discontinuities of surface orientation, and reflectance edges. Each edge type imposes distinct constraints on surface recovery, but these constraints cannot be exploited unless edges can be reliably classified. Existing edge-classification techniques based on junction catalogues and constraint propagation depend critically on ideal data, and are therefore inadequate for natural imagery. For these reasons, our work focused on developing new edge-classification techniques that could be applied to natural imagery, and as a result of this effort, we developed and implemented a new intensity-based approach to edge classification. Using basic properties of scenes and images, we deduced signatures for several edge types that are expressed in terms of correlational properties of the image intensities in the neighborhood of the edge, and developed a computer program that evaluates image edges against these prototype signatures. The program effectively discriminates extremal boundaries from cast shadow boundaries in cases where traditional junction cues are absent from the image.

Reports of our previous work on edge reconstruction, surface interpolation, and shape recovery from texture have been published in professional journals [Ref 1-7]; reprints of these papers are available on request.

### A. Edge Classification

Edges play a central role in three-dimensional surface reconstruction. Crucial to exploiting the constraints imposed by edges is edge sorting--classifying the image edges according to the type of surface boundary they represent (e.g., extremal boundaries, shadow edges, surface orientation discontinuities, or texture edges). Because each edge type imposes different constraints on three-dimensional

interpretation, misclassification can lead to serious interpretation errors. Edge classification in line drawings has been addressed in terms of propagation of junction constraints, global structural cues such as parallelism and symmetry, and global optimization criteria on the three-dimensional interpretation. Because these techniques depend on perfect edge data, their applicability to natural imagery is questionable.

An alternative approach to edge sorting is to use intensity and spectral information in the neighborhood of the edge. Horn [8] suggested that the intensity profiles across edges (such as peak versus step) could provide signatures for some edge types. However, this technique has not worked for complex imagery.

In this section we describe an intensity-based, line-sorting technique that distinguishes line types by statistically comparing intensity variations along opposite sides of the edge. We have focused on two line types--extremal edges and cast shadow boundaries--but extensions to other edge types have also been explored.

#### B. Defining the Problem

Because line types are defined in terms of the scene events they denote, any method for line sorting must provide some basis for discriminating those events by their appearance in the image. We therefore begin by characterizing the distinctive properties of extremal boundaries and cast shadow edges, and defining the computational problem of identifying those edges.

Extremal Boundaries--Projective mapping from image to scene tends to be continuous because physical surfaces tend to be continuous. Almost everywhere in a typical image, therefore, nearby points in the image correspond to nearby points in the scene. This adjacency is preserved over any change in point of view or scene configuration, short of rending the connected surfaces of which the



scene is composed. The distinguishing property of extremal boundaries (which can be defined as discontinuities in the projective mapping) is their systematic violation of this rule: the apparent juxtaposition of two surfaces across an extremal edge represents no fixed property of either surface, but is subject to the vagaries of viewpoint and scene configuration. For example, if you position your finger to coincide with a particular feature on the wall or outside the window, a small change in the position of head or hand may drastically affect the apparent relationship. Because the false appearance of proximity is the hallmark of extremal edges, the problem in identifying those edges is to distinguish in the image the actual proximity of nearby points on connected surfaces from accidental proximity imposed by projection.

Cast Shadows--Cast shadows in outdoor scenes represent transitions from direct to scattered illumination caused by the interposition of an occluding body between the sun and the viewed surface. The problem in identifying cast shadows is to distinguish these transitions in incident illumination from changes in albedo or surface orientation, for example. This kind of discrimination presents a problem because the effects of all these parameters are confounded in the image data--a change in image brightness may reflect a change in albedo or surface orientation, as well as in incident illumination. Because the relation among illumination, reflectivity, orientation, and image irradiance is well known, the presence of shadows in an image could be readily detected if a constant planar reference pattern could be placed in the scene; when the apparent brightness of a constant pattern varies with location, the change in brightness must, by elimination, be attributed to a change in

illumination. Of course, such active intervention is generally impractical; the problem is to achieve the effect of viewing a constant pattern across the shadow edge without actually placing such a pattern in the scene. This effect could be achieved if some fixed relationship were known to hold between the surface strips on each side of the shadow edge.

In short, extremal boundaries are curves across which distant points in space are placed in apparent juxtaposition by projection, violating the continuity of the projective mapping that holds over most of the image. To identify extremal boundaries requires, therefore, that actual proximity be distinguished from apparent proximity imposed by projection. Cast shadow edges are contours across which the pattern of surface reflectance has been systematically transformed by an abrupt change in illumination. To identify cast shadow edges, the effects of illumination must be distinguished from those of albedo and surface orientation, as if a constant planar reference pattern had been placed across the edge.

### C. Computational Theory

Our solution rests on the simple principle that coherence in the image reflects real coherence in the scene, rather than a coincidence of the structure and alignment of distinct scene constituents. We measure coherence in the neighborhood of an edge by performing a normalized correlation on intensity values at corresponding points across the edge. (Other measures of coherence are possible, such as continuity of linear structure.)

A high correlation implies that the edge and its neighborhood correspond to a strip on a connected surface. Therefore, the edge is not an extremal boundary, and furthermore, the regions on either side can be regarded as instances of a (statistically) constant pattern. In that case, the presence of a shadow can be detected by constructing a regression equation whose parameters signal any systematic photometric

distortion of the pattern across the edge. Ideally, this distortion is linear, but nonlinearities are introduced in practice by complex lighting effects, film or sensor response, and so forth.

A low correlation does not necessarily signal an extremal boundary, but could reflect low contrast or fragmented surface structure. However, the local disruptions of correlation that signal extremal edges can be distinguished from a global lack of structured surface markings by using a neighborhood of the image around the edge to set a baseline for correlation. A contour of low correlation surrounded by regions of high correlation is likely to denote an extremal boundary.

To obtain a baseline, the given edge is embedded in a family of parallel curves, and a sequence of regressions performed from one curve onto the next. In terms of this regression sequence, the various edge types display distinctive "signatures" that can be computed from the image data: extremal boundaries display a sharp notch in correlation where the fabric of the projective mapping is torn by the boundary. Cast shadow boundaries sustain high correlations across the edge, but sharp spikes occur in the regression parameters where the surface structure is systematically transformed by the illumination transition. A low correlation throughout implies that either the contrast is too low or the surface structure too fragmented for any positive conclusion to be drawn.

This strategy follows from the assumption that coherence in the image--as measured by correlation--implies a connected surface. The rationale for this assumption follows from some elementary observations on the character of natural scenes and images. First, as mentioned above, it follows from the fact that surfaces tend to be continuous so that nearby points in the image usually correspond to nearby points in the scene (i.e., the projective mapping is, in general, continuous). Second, because the structure of surfaces tends to be coherent, such properties as reflectance and orientation at a given point on a connected surface are (statistically) good predictors of the properties at nearby points. Third, because scenes are made up of distinct objects

whose structure and spatial configuration are governed by extremely complex factors, the properties of widely separated surface points, or points on surfaces of distinct objects, can usually be regarded as unrelated and independent.

Because of these three principles--surface continuity, coherence, and independence--we can expect intensity values at nearby image points to be highly correlated. (It is easily verified that this is so for most images.) That is, a small step in the image usually corresponds to a small step on some connected surface, so surface coherence imposes a statistical relation on the properties of nearby points. Thus, when we place the points on either side of an arbitrary image curve in correspondence, we should often expect to see a high correlation between the intensity values at those points. However, when that small step happens to cross an extremal boundary, the corresponding surface points, belong in general, to distinct objects, and might be widely separated in space. In that case, the properties of the points are independent. Thus, when the points on either side of an extremal boundary are placed in correspondence, we should never observe a high correlation unless the surfaces meeting at the boundary possess identical structures, and happen to lie in perfect register from the observer's viewpoint. The likelihood of this occurrence is vanishingly small.

Thus, we may confidently conclude that coherence of structure across an image curve (as measured by correlation) denotes true coherence of scene structure rather than an accident of scene configuration.

#### D. Implementation and Results

Our implementation assumes that an edge has been located by edge-finding techniques. In practice, edges were often traced by hand; automatically detected zero-crossing edges were also used as inputs. We construct a parallel family of curves around the edge by imposing a new coordinate system on the image as follows: arc length on the edge is taken as the y-coordinate, and orthogonal distance from the edge (right-

handed) as the x-coordinate. This amounts to coercing a strip around the edge into a rectangular region whose central column corresponds to the original edge. The surrounding columns correspond to parallel curves on either side of the edge. The rectangular strip was constructed using bilinear interpolation of intensity values to reduce quantization artifacts.

Once the rectified strip was constructed, a sequence of linear regressions was performed between columns. To avoid spurious correlation imposed by the imaging and digitizing process, regressions were computed between the  $i$ -th column and the  $(i + 2)$ th. The outcome of this computation was a normalized correlation coefficient, additive regression term, and multiplicative regression term, each a function of the location of the column. The midpoints of these plots represent the regression across the original edge.

In terms of these regression sequence plots, we define the following expected edge signatures (see Figure 1 for Idealized plots):

- \* An extremal boundary is indicated by a sharp notch in an otherwise high correlation at the nominal edge location.
- \* A cast shadow boundary is indicated by high correlation maintained across the edge, but sharp spikes or notches can be present in the additive and multiplicative regression parameter, depending on the sense of the shadow transition and the digitization function.
- \* A high correlation coefficient with no disturbance in the regression parameters implies that the edge is not physically significant.
- \* Sustained low correlation implies low contrast or lack of surface structure, and no classification can be made.

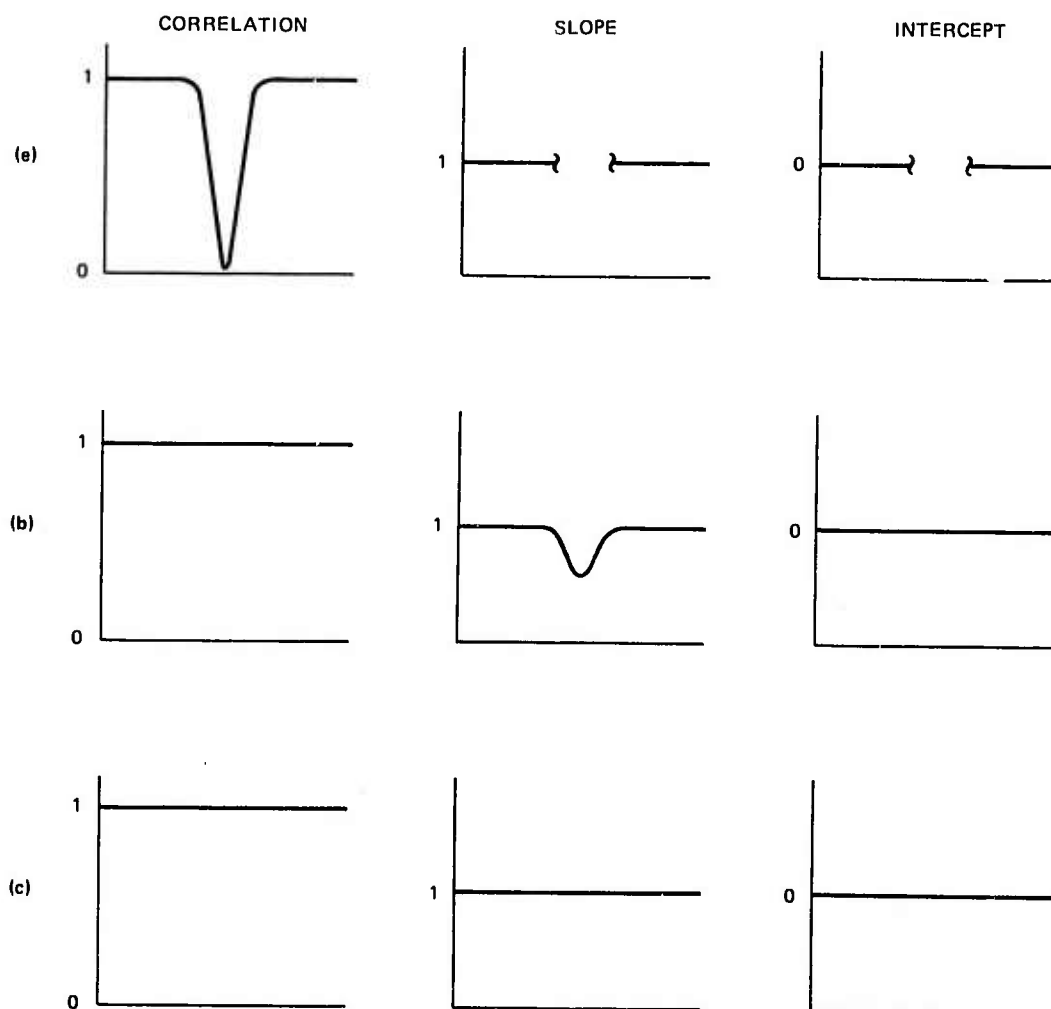
No attempt has yet been made to classify the edge type signatures automatically; however, the computation was performed on a number of edges in both aerial and ground imagery. Examples of the images, edges, and regression sequence plots are shown in Figures 2 through 6. The regression plots should be compared to the idealized signatures of Figure 1.

The edge-sorting method presented above, derived from basic properties of visual scenes, shows promise as a useful technique, particularly in connection with established line-junction techniques. Potential specialized applications of the technique include shadow detection for use in raised-object cueing and camera-model recovery.

A detailed report of this technique is being prepared for publication.

#### REFERENCES

1. A. P. Witkin, "Recovering Surface Shape and Orientation from Texture," Artificial Intelligence Journal, Volume 17 (1-3), pp. 17-45 (August 1981).
2. H. G. Barrow and J. M. Tenenbaum, "Interpreting Line Drawings as Three-Dimensional Surfaces," presented at First Annual AAAI Conference, Stanford University, Stanford, California (August 1980); also in Artificial Intelligence Journal, Volume 17 (1-3) (August 1981).
3. J. M. Tenenbaum, M. A. Fischler, and H. G. Barrow, "Scene Modeling: A Structural Basis for image Description," Computer Graphics and Image Processing, Volume 12(4), pp. 407-425 (April 1980); also in Image Modeling, A. Rosenfeld, ed., pp. 371-389 (Academic Press, New York, New York, 1981).
4. H. G. Barrow and J. M. Tenenbaum, "Computational Vision," Proceedings IEEE, Volume 69(5), pp. 572-595 (May 1981).
5. A. P. Witkin, "A Statistical Technique for Recovering Surface Orientation from Texture in Natural imagery," presented at First Annual AAAI Conference, Stanford University, Stanford, California (August 1980).
6. H. G. Barrow and J. M. Tenenbaum, "Reconstructing Smooth Surfaces from Partial, Noisy information," Proceedings of the ARPA IU Workshop, University of Southern California, Los Angeles, California (November 1979).
7. H. G. Barrow and J. M. Tenenbaum, "Recovering intrinsic Scene Characteristics from images," in Computer Vision Systems, A. Hanson and E. Riseman, eds., pp. 3-26 (Academic Press, New York, New York, 1978).
8. B. K. P. Horn, "Understanding Image Intensities," Artificial Intelligence, Volume 8, No. 2, pp. 201-231 (1977).



- (a) **Extremal Boundary** — notch in correlation across the edge. Slope and Intercept in the low-correlation area are meaningless.
- (b) **Cast Shadow** — sustained high correlation across the edge, with disturbance of one or both regression parameters. The nature of this disturbance depends on the sense of the edge (i.e. whether the shadow lies on the left or right), and on details of the imaging and digitizing process. In practice, nonlinearities perturb the correlation slightly.
- (c) **No Edge Present** — sustained high correlation, no disturbance in regression parameters.

FIGURE 1 IDEALIZED REGRESSION PLOTS

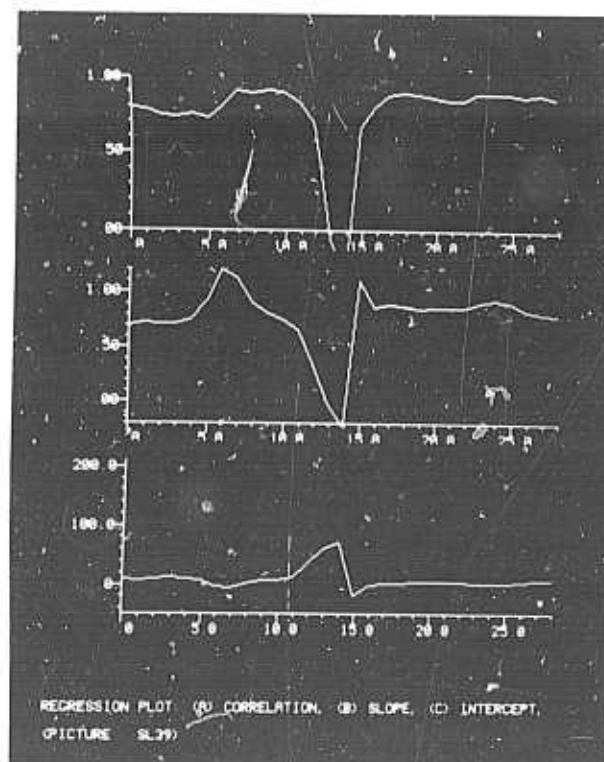
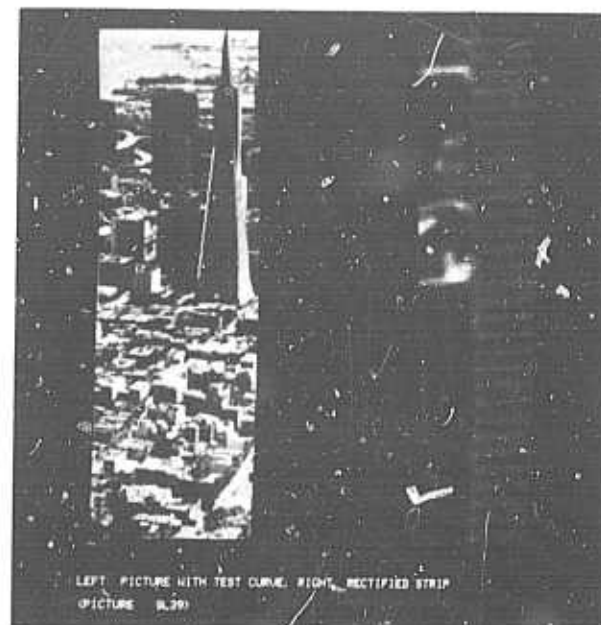


FIGURE 2 EXAMPLE OF EXTREMAL EDGE



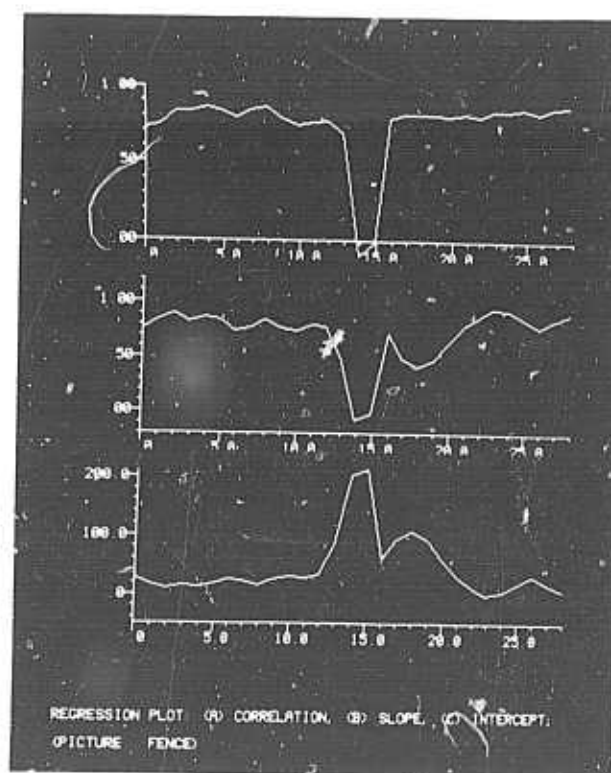
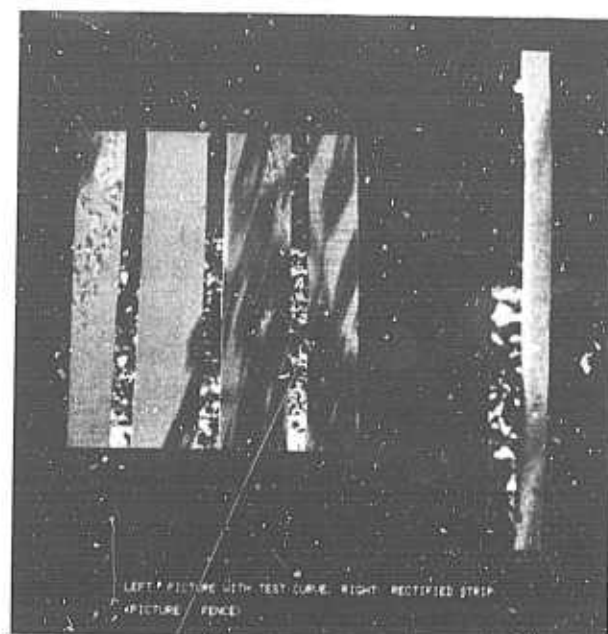


FIGURE 3 EXAMPLE OF EXTREMAL EDGE

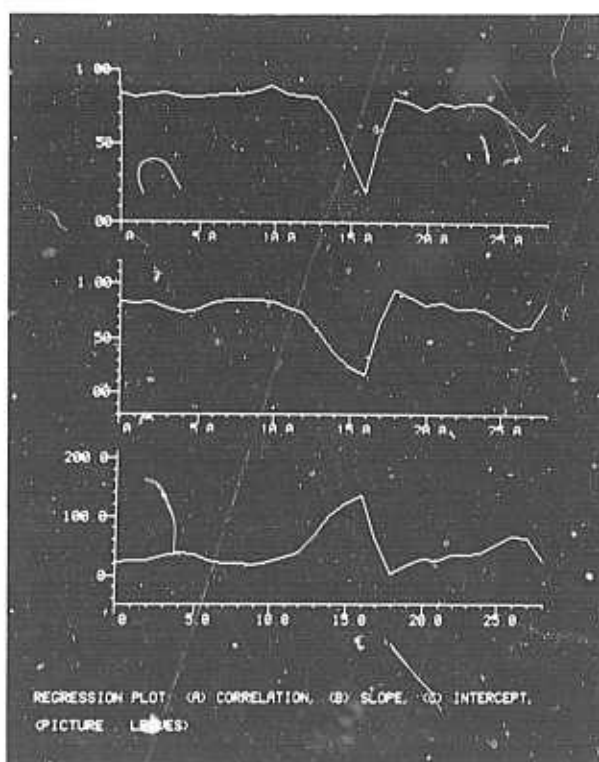
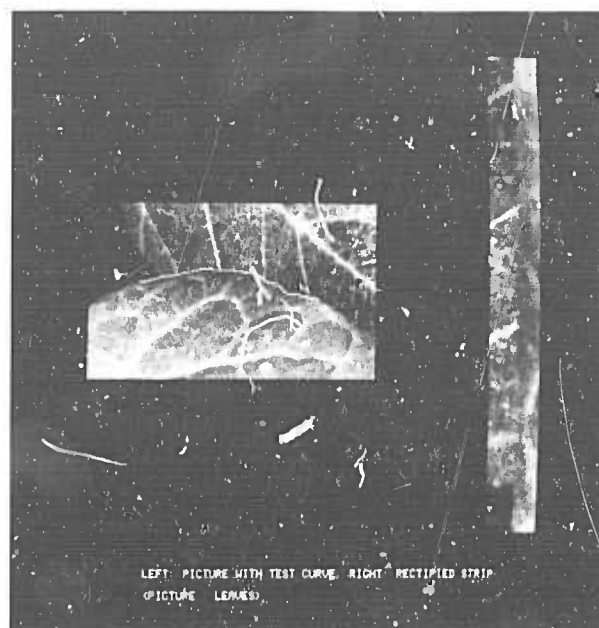


FIGURE 4 EXAMPLE OF LOW-CONTRAST EXTREMAL EDGE

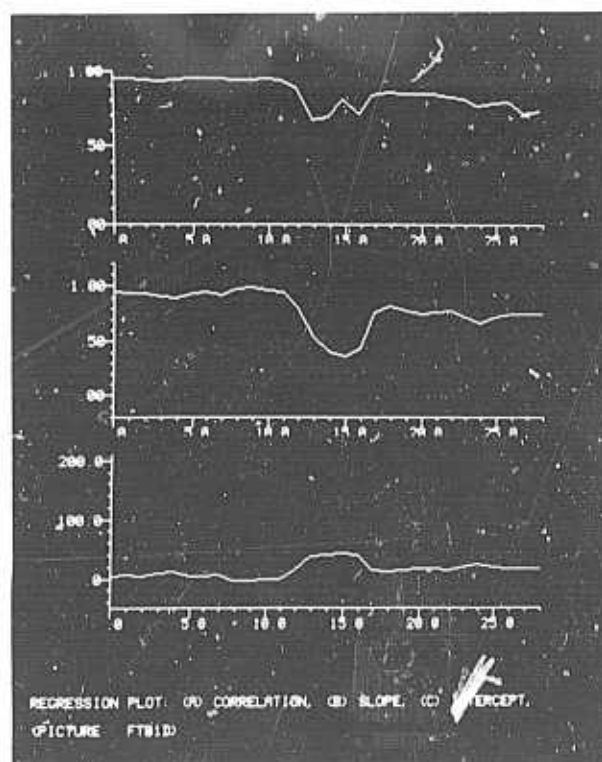
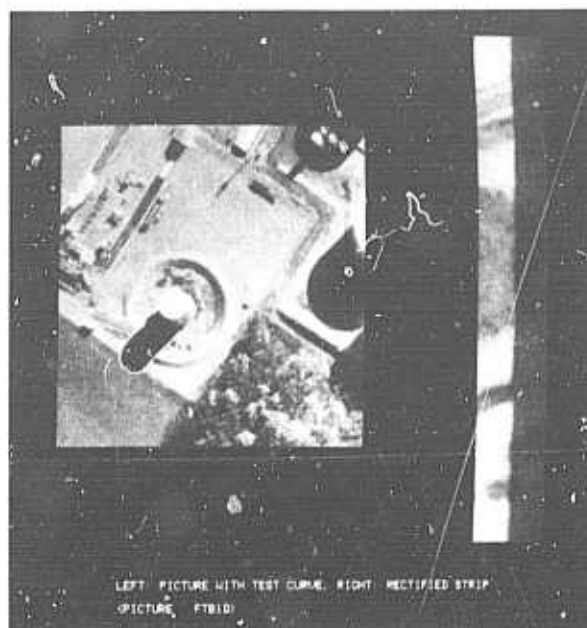


FIGURE 5 EXAMPLE OF CAST-SHADOW EDGE

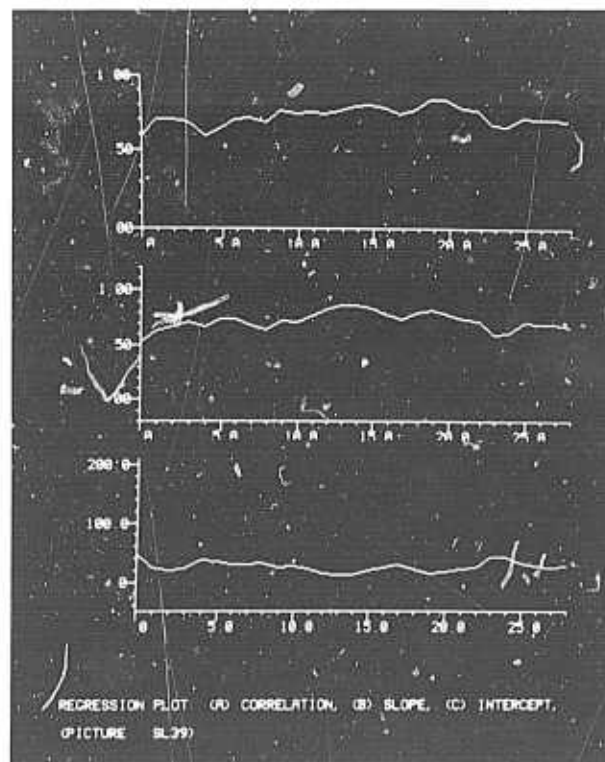
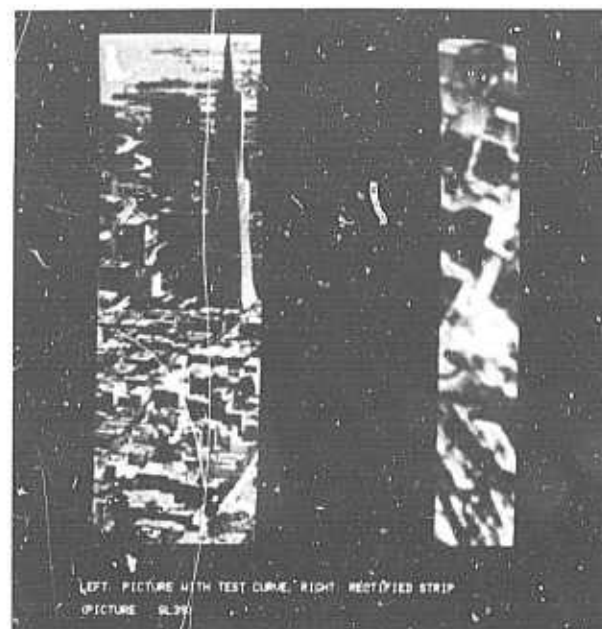


FIGURE 6 EXAMPLE OF REGRESSION PLOTS WHERE NO EDGE IS PRESENT